

Deep-learning-assisted microscopy with ultraviolet surface excitation for rapid slide-free histological imaging

ZHENGHUI CHEN, WENTAO YU, IVY H. M. WONG, AND TERENCE T. W. WONG*

Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

*ttwwong@ust.hk

Abstract: Histopathological examination of tissue sections is the gold standard for disease diagnosis. However, the conventional histopathology workflow requires lengthy and laborious sample preparation to obtain thin tissue slices, causing about a one-week delay to generate an accurate diagnostic report. Recently, microscopy with ultraviolet surface excitation (MUSE), a rapid and slide-free imaging technique, has been developed to image fresh and thick tissues with specific molecular contrast. Here, we propose to apply an unsupervised generative adversarial network framework to translate colorful MUSE images into Deep-MUSE images that highly resemble hematoxylin and eosin staining, allowing easy adaptation by pathologists. By eliminating the needs of all sample processing steps (except staining), a MUSE image with subcellular resolution for a typical brain biopsy (5 mm × 5 mm) can be acquired in 5 minutes, which is further translated into a Deep-MUSE image in 40 seconds, simplifying the standard histopathology workflow dramatically and providing histological images intraoperatively.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Histopathological examination is essential for tissue-based disease diagnosis. The gold standard is to obtain histologically stained images of thin tissue sections under a brightfield microscope, which requires lengthy and laborious tissue preparation steps: (i) the tissue sample should be formalin-fixed and paraffin-embedded (FFPE), (ii) sectioned by a microtome to get optically thin slices (4–7 μm), (iii) picked up onto slides in the floatation bath and deparaffinized (iv) stained with histological stains (e.g., hematoxylin and eosin (H&E)), and (v) mounted on a glass slide (Fig. 1(A)). The entire procedure takes at least a day, and usually requires 3–5 days [1], causing a delay for the disease diagnosis, thus, imposing an additional physical and mental burden on patients. The alternative workflow for intraoperative histology can be achieved by frozen section, which takes ~20–30 mins to process the specimen (including freezing, cutting, and staining). However, the image quality of the frozen section is often unsatisfactory compared to standard histology provided by the FFPE tissue section due to the artifacts caused by poor sectioning and staining qualities [2]. To address this pressing clinical need, it is important to develop a new approach that could provide fast and slide-free histological imaging with high image quality and molecular specificity for accurate intraoperative assessment.

In recent years, various slide-free imaging modalities have been proposed to eliminate the need for a lengthy sample preparation procedure. These approaches include confocal microscopy [3], photoacoustic microscopy [4,5], nonlinear microscopy [6–8], structured illumination microscopy [9], and light-sheet microscopy [10]. To improve the interpretability of the images acquired by these slide-free imaging techniques, different models have been utilized to transform these images into histologically stained images, which can be divided into two categories, including analytical-based and deep-learning-based approaches. For the analytical-based approach, the

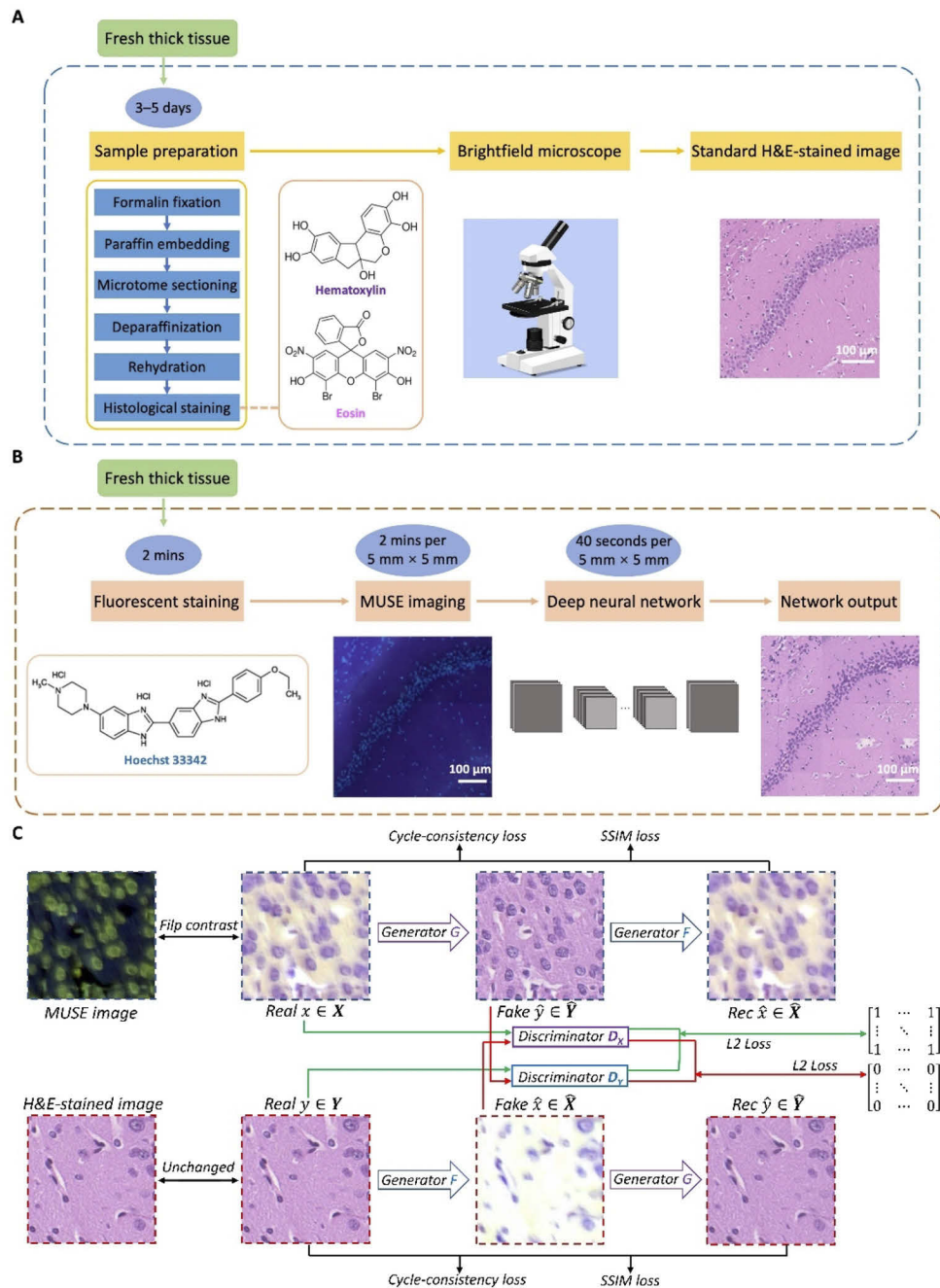


Fig. 1. Illustration of the histopathology workflow without and with Deep-MUSE images. (A) A standard histopathology workflow for whole-slide imaging. (B) The proposed histopathology workflow with Deep-MUSE images. The deep neural network generates a Deep-MUSE image that is comparable to the conventional H&E-stained image after inputting a MUSE image acquired from fresh mouse brain tissue, bypassing the complex sample preparation procedures. (C) The basic structure of CycleGAN to learn the mapping function between the MUSE image (e.g., FFPE mouse brain slices) and any H&E-stained image. Two generator-discriminator pairs are used for model training using unpaired image data. Rec: recovered; SSIM: structural similarity.

virtual H&E-stained images are rendered from the slide-free grayscale images of the labeled sample either by a simple linear color-coded model [11] or nonlinear absorption model using Beer-Lambert-law [12]. These two models require prior knowledge about the property of the fluorescent dyes in the biological tissues, which is normally hard to predict in practice, hindering their reliability in generating high-quality virtual H&E-stained images.

For the deep-learning-based approach, the data-driven model employs the deep convolutional neural network (DCNN) to transform label-free images into virtual H&E-stained images for FFPE thin slices. The generative adversarial network (GAN) [13] is an advanced framework typically consisting of two DCNN models for adversarial training, generating high-quality virtual H&E-stained histological images from different imaging modalities such as autofluorescence imaging [14] and quantitative phase imaging [15]. A type of GAN model learns the style mapping function from any label-free images to brightfield H&E-stained images through paired image data training. This conditional GAN with a supervised concept is known as pix2pix [16], achieving image-to-image translation with rigidly paired label-free images and the corresponding H&E-stained images of the same FFPE thin slices for training. As the ground truth H&E-stained images can only be obtained for thin sections, the pix2pix model can only be applied to FFPE thin tissue sections. In addition, the image co-registration process between different image modalities is complicated, including the coarse global registration and accurate local registration, which takes much effort for image data alignment [2]. The precision level of image pairs will also be affected easily by other external conditions. For instance, the cellular information will be slightly changed when FFPE tissue slides go through the H&E staining process, causing an inherent difference between the two sets of images. To get rid of the image registration procedure, unsupervised GAN frameworks such as UNIT [17] and CycleGAN [18] are developed to perform style transfer between two different image domains, e.g., Fourier ptychographic microscopy (FPM) images and immunohistochemistry (IHC) staining images [19].

Unsupervised deep learning models have been applied to nonlinear multimodal imaging [20] and confocal microscopy [21] to obtain virtual H&E-stained images for thick specimens. However, these imaging systems require expensive lasers and point scanning, hindering their wide adaptation in hospitals. Microscopy with ultraviolet surface excitation (MUSE) is a novel slide-free imaging technique with a simple design and low cost, which can be an alternative for histological examination after an H&E conversion by different image translation models [22]. In this regard, here, we propose an improved unsupervised deep-learning-based model with the integration of microscopy with ultraviolet surface excitation (Deep-MUSE). Microscopy with ultraviolet surface excitation (MUSE) is a rapidly emerging fluorescence-based slide-free imaging modality with the advantage of simple, fast, and cost-effective [23]. The ultraviolet light-emitting diode (UV-LED) excites the labeled cell nuclei, providing color-contrast that is equivalent to the conventional histology [24]. The penetrating depth of UV light is merely around several microns similar to the thickness of the common histology slice, allowing the fluorescence excitation to be confined only on the surface of a tissue sample [25]. This key feature eliminates the need for microtome sectioning of the specimen, bypassing the laborious sample preparation steps. In the previous study [23], virtual staining was performed based on Beer-Lambert physical model of transillumination microscopy, which requires prior knowledge about the RGB color coordinates of fluorescent dyes and requires region selection by users to correct the presence of spectrally mixed signals. Here, we propose that the H&E staining can be performed without any manual tuning by the inference of the feed-forward network after the unpaired training for the CycleGAN model has been completed. We demonstrate that our Deep-MUSE images are comparable to standard H&E-stained images on mouse brain tissue with various sample preparation protocols, including FFPE thin slices and fixed/fresh thick samples. Hence, this approach is highly suitable for rapid histopathological diagnosis, which is a captivating alternative to the gold standard of intraoperative assessment — frozen section.

2. Materials and methods

2.1. Sample preparation and staining

In this study, FFPE thin slices, formalin-fixed and thick, fresh and thick mouse brain, and formalin-fixed thick mouse liver samples were used. Mice were supplied from the Animal and Plant Care Facility at the Hong Kong University of Science and Technology (HKUST). The experiments were performed with the approval of the Animal Ethics Committee and the medical surveillance of the Health, Safety & Environment Office at HKUST.

For FFPE sample preparation (Fig. S1), the harvested mouse brain was fixed with 4% formalin solution for one day. Then, the fixed mouse brain was embedded in a paraffin block, cut into thin sections (7- μm thick) by a microtome (HistoCore AUTOCUT, Leica Biosystems), and placed on a quartz slide. After deparaffinization of the tissue slice with xylene, Hoechst 33342 (500 $\mu\text{g ml}^{-1}$ in phosphate-buffered saline (PBS)) solution was used to stain the sample for about 20 s. After imaging by the MUSE system, the same slice was stained with H&E following the conventional protocol to get the ground truth histological image.

For fixed and thick mouse brain sample preparation, after embedding the fixed mouse brain into 2% agarose gel, vibratome was used to shave off several 100- μm thick slices until the remaining thick brain sample has a flat surface. Afterward, we used the remaining thick brain with a thickness of several millimeters for the experiment. A low concentration Hoechst 33342 (50 $\mu\text{g ml}^{-1}$ in PBS) solution was used to stain the sample for about 2–3 mins, which ensures only the tissue surface was labeled. After imaging with the MUSE system, the thick sample was paraffin-embedded, sectioned for a series of layers adjacent to the surface of the block, and stained with H&E to get the reference histological image. For fixed and thick mouse liver sample preparation, the whole procedure is the same except mouse liver sample was stained with DAPI (5 $\mu\text{g ml}^{-1}$ in PBS) solution.

For fresh and thick mouse brain sample preparation, we cut the brain into halves using a typical blade by hand. Then, we stained the brain with a low concentration Hoechst 33342 (50 $\mu\text{g ml}^{-1}$ in PBS) solution for about 2–3 mins before MUSE imaging. After that, we put the stained brain into a 4% formalin solution for fixation, embedded it in a paraffin block, and sectioned it into thin slices (7- μm thick). Finally, we acquired the reference bright-field histological image with H&E staining. All the above H&E histological images were obtained with a digital slide scanner (NanoZoomer, Hamamatsu Photonics K.K.).

2.2. Data acquisition and processing

The slide-free MUSE images were captured using our lab-built inverted fluorescence microscope with a motorized stage (Fig. S2). The specimen was put on a quartz slide which is UV transparent, and the slide was placed on the sample holder connected to a three-axis motorized stage. A UV light from an obliquely oriented 265-nm LED (M265L4, Thorlabs Inc.) was focused on the Hoechst-stained sample. Subsequently, the excited fluorescence signals were collected using a 10 \times objective lens (RMS10X-PF, NA = 0.3, Thorlabs Inc.), and focused on a color camera (DS-Fi3, 5.9 megapixels, Nikon Inc.) using an infinity-corrected tube lens (TTL180-A, F = 180 mm, Thorlabs Inc.).

We developed a LabVIEW-based software user interface to acquire data by delivering instructions to a single-chip microcomputer (SCM, Mega 2560, Arduino) in order to trigger the camera and motorized stage to scan different areas along x and y directions. The exposure time of the camera for each FOV ($\sim 691 \mu\text{m} \times 491 \mu\text{m}$) is ~ 500 ms. The mosaic FOVs would be stitched with a 20% overlapping area in MATLAB after scanning the whole slice.

For FFPE samples, the MUSE and the corresponding H&E-stained images of the same thin section can be provided. Therefore, it is possible to implement both pix2pix and CycleGAN

models, and then evaluate their performance by comparing them with the ground truth H&E-stained images. The ROI between these two imaging modalities was registered in MATLAB through a complex image registration process. First, the global registration process scaled the pixel size, achieving rotation and shift corrections between the same ROI of the MUSE and H&E-stained images [26]. Then, the local registration process was performed to accurately register the input and target image pairs at the patch level by applying an elastic image registration algorithm [27]. These image pairs were used for pix2pix model training and evaluation. The size of the image patch was 256×256 . The training datasets include 1011 paired MUSE and H&E-stained image patches from the same FFPE mouse brain slice. The validation datasets include 479 paired MUSE and H&E-stained image patches from another FFPE mouse brain slice. For the CycleGAN model, the unpaired training datasets include 5883 MUSE and 5846 H&E-stained image patches.

For fixed mouse brain samples, the CycleGAN model was trained with unpaired 1505 MUSE and 1502 H&E-stained image patches. The validation datasets include 3600 MUSE image patches from another fixed mouse brain sample. For fixed mouse liver samples, the CycleGAN model was trained with unpaired 1288 MUSE and 1481 H&E-stained image patches. The validation datasets include 805 MUSE image patches. For fresh and thick mouse brain samples, the CycleGAN model was trained with unpaired 3177 MUSE and 2574 H&E-stained image patches. The validation datasets include 2500 MUSE image patches from another fresh mouse brain sample.

2.3. Network architecture

The GAN [13] architecture is applied to learn the style transformation from slide-free MUSE images to their equivalent bright-field H&E-stained images. A standard DCNN model uses an encoder-decoder network [28] for image translation tasks. For example, the popular encoder-decoder model “U-Net” [29] adds skip connections to transfer the low-level information directly across the network. The $L1$ loss function (mean absolute error) or $L2$ loss function (mean squared error) are the general option to minimize the distance between network output image and target image. However, these two loss functions tend to produce blurry images caused by the pixel-wise average of plausible results in the pixel space [30]. To address this problem, GANs can output sharp and realistic images through the adversarial training of two DCNN models, including a generative model (Generator) and a discriminative model (Discriminator). The generator G aims to fool the discriminator D by producing fake images similar to target images, while D intends to classify the generated fake image and real target image. In our case, the generator G learns the statistical transformation between MUSE and H&E-stained images, whereas the discriminator D aims to distinguish Deep-MUSE images (i.e., fake generator output) from real bright-field H&E-stained images. We used two types of GAN models, including pix2pix [16] for paired image translation tasks and CycleGAN [18] for unpaired image translation tasks.

Taking the FFPE samples as a demonstration, we can use these two GAN models. The architecture of pix2pix (Fig. S3(A)) includes one generator G and one discriminator D . The generator G translates image domain X (MUSE images) into image domain Y (H&E-stained images). The MUSE image x is fed into the generator G and output the fake virtual H&E-stained image (Deep-MUSE image) $\hat{y} = G(x)$. The discriminator D makes a judgment about the concatenation of the MUSE image x with the fake virtual H&E-stained image \hat{y} or real H&E-stained image y . The loss function of GAN can be expressed as:

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{x,y}[D(x, y)^2] + \mathbb{E}_x[(1 - D(x, G(x)))^2] \quad (1)$$

The output fake virtual H&E-stained image \hat{y} should also be similar to the ground truth H&E-stained image y . We used the $L1$ loss function to measure the distance between these paired

images:

$$\mathcal{L}_1(G) = E_{x,y}[||y - G(x)||_1] \quad (2)$$

The full objective of pix2pix is:

$$G^* = \arg \min_G \max_D \mathcal{L}_{GAN}(G, D) + \lambda \mathcal{L}_1(G) \quad (3)$$

For many applications, it is hard to acquire paired data for pix2pix model training. To overcome this challenge, the architecture of CycleGAN (Fig. 1(C)) was designed for performing image translation with unpaired data. From the preliminary experiment, we found that the model trained by original MUSE images and H&E-stained images will output low-quality Deep-MUSE images (Fig. S4). The model converted bright cell nuclei and dark background in MUSE images into a white background and purple nuclei in Deep-MUSE images, respectively. The general tendency is that the CycleGAN model will map the dark color (black) in MUSE images into the dark color (purple) in H&E-stained images. To address this problem, we have inverted the color of MUSE images for model training and inference, which leads to an accurate style transfer function.

We assigned the inverted MUSE image x to image domain X , and the H&E-stained image y to image domain Y . The generator G maps image domain X to image domain Y by transforming the real MUSE image x into fake virtual H&E-stained image $\hat{y} = G(x)$. The generator F maps image domain Y to image domain X by transforming real H&E-stained image y into fake MUSE image $\hat{x} = F(y)$. The discriminator D_X classifies real MUSE image x and fake MUSE image \hat{x} . The discriminator D_Y classifies real H&E-stained image y and fake virtual H&E-stained image \hat{y} . The objective of the adversarial loss function is:

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{x \sim p_{data}(x)}[(D_Y(G(x)) - 1)^2] + \mathbb{E}_{y \sim p_{data}(y)}[D_Y(y)^2] \quad (4)$$

$$\mathcal{L}_{GAN}(F, D_X, Y, X) = \mathbb{E}_{y \sim p_{data}(y)}[(D_X(F(y)) - 1)^2] + \mathbb{E}_{x \sim p_{data}(x)}[D_X(x)^2] \quad (5)$$

With this GAN loss only, it is not possible to guarantee successful transformation between the MUSE image and H&E-stained image without pair images for supervised training. It is important to make sure that all the content in MUSE images can be preserved (e.g., the nuclear size and count), and the style should be visually similar to H&E-stained images during the transformation. By combining the mapping function of the generators G and F , we can impose an extra constraint named cycle consistency loss [18]. A MUSE image x from domain X will become recovered image $\hat{x} = F(G(x))$ after passing the generators G and F , i.e., $F(G(x)) = x$. Similarly, an H&E-stained image y from domain Y is converted to recovered image $\hat{y} = G(F(y))$ after going through the generators F and G , i.e., $G(F(y)) = y$. Therefore, we used L1 loss between original images and recovered images as this cycle-consistency loss:

$$\mathcal{L}_{cyc}(G, F) = \mathbb{E}_{x \sim p_{data}(x)}[||F(G(x)) - x||_1] + \mathbb{E}_{y \sim p_{data}(y)}[||G(F(y)) - y||_1] \quad (6)$$

To further improve the perceptual quality of the recovered images, we also calculated the structural similarity (SSIM) index [31] between the original image x and recovered image y as the perceptual distance. The SSIM index is composed of three functions including luminance $l(x, y)$, contrast $c(x, y)$, and structure $s(x, y)$:

$$SSIM(x, y) = [l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma] \quad (7)$$

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (8)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (9)$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \quad (10)$$

where μ_x and μ_y refer to the mean values of image x and image y , respectively. σ_x and σ_y refer to the standard deviations of image x and image y , respectively. σ_{xy} stands for the covariance between image x and image y , and c_1 and c_2 stand for constants. The expression can be simplified by setting $\alpha = 1, \beta = 1, \gamma = 1$ and $c_3 = c_2/2$. The value of the SSIM index is ranged from zero to one. Then, the loss function is constructed by one minus the SSIM index. In our experiment, the CycleGAN model trained with SSIM loss outputs smooth and authentic Deep-MUSE images (Fig. S5). SSIM loss is calculated between original images x, y , and the recovered images $\hat{x} = F(G(x)), \hat{y} = G(F(y))$:

$$\mathcal{L}_{ssim}(G, F) = \mathbb{E}_{x \sim p_{data}(x)}[||F(G(x)) - x||_{ssim}] + \mathbb{E}_{y \sim p_{data}(y)}[||G(F(y)) - y||_{ssim}] \quad (11)$$

The generators G and F should be identity mapping for input H&E-stained image y and MUSE image x , respectively [32], i.e., $G(y) = y$ and $F(x) = x$. The $L1$ loss function can be used to measure the image distance:

$$\mathcal{L}_{idt}(G, F) = \mathbb{E}_{y \sim p_{data}(y)}[||G(y) - y||_1] + \mathbb{E}_{x \sim p_{data}(x)}[||F(x) - x||_1] \quad (12)$$

The final objective of CycleGAN becomes:

$$\begin{aligned} G^*, F^* = \arg \min_{G, F} \max_{D_X, D_Y} & \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(F, D_X, Y, X) \\ & + \lambda_1 \mathcal{L}_{cyc}(G, F) + \lambda_2 \mathcal{L}_{idt}(G, F) + \lambda_3 \mathcal{L}_{ssim}(G, F) \end{aligned} \quad (13)$$

The generator and discriminator architectures (Fig. S3(B)) are the same in pix2pix and CycleGAN. Only the input images of the generator and discriminator are different as mentioned above. The generator is adopted from the image transformation network that contains convolutional layers for image downsampling and upsampling [33]. Several residual blocks [34] connect the downsampling and upsampling blocks. The discriminator networks use PatchGAN [16] to preserve the high-frequency structures in the images. The real or fake images will be classified at the local image patch level after passing through the convolutional layers.

For pix2pix experiments, we set $\lambda = 10$ in Eq. (3) and train the model with 155 epochs. For CycleGAN experiments, we set $\lambda_1 = 10, \lambda_2 = 5, \lambda_3 = 1$ in Eq. (13) and train the model with 70 epochs. The networks were trained from scratch with a learning rate of 0.0002. The Gaussian distribution with a mean = 0 and standard deviation = 0.02 was used to initialize the weights of the network. An Adam optimizer was used to minimize the overall loss functions.

2.4. Novel workflow

By the integration of the MUSE system and deep neural network, we enable a new histopathology workflow (Fig. 1(B)) that can provide rapid slide-free histological imaging of unprocessed specimens with various thicknesses. The MUSE system (Fig. S2) is applied to generate fluorescence images of a fresh tissue specimen by UV-LED illumination directly. This system can achieve high resolution and high imaging contrast with the visible light emission of fluorescent dyes excited by UV light, which is sufficient to reveal the morphology of individual cell nuclei. The MUSE image quality is comparable to the bright-field H&E-stained images obtained from a digital slide scanner using an objective lens with the same numerical aperture (NA). The unsupervised deep learning model (Fig. 1(C)) learns the style transfer function between the MUSE and H&E-stained images after the bidirectional training of unpaired image data. In whole-slide imaging, we have considered an imaging FOV of $5 \text{ mm} \times 5 \text{ mm}$, which is a typical size of a brain biopsy for comparison. The MUSE imaging system takes ~6 mins only to generate Deep-MUSE images for the entire workflow, whereas the NanoZoomer takes ~2–3 mins to obtain

a brightfield image with the same FOV after lengthy sample preparation steps. We validate the proposed workflow using mouse brain tissue samples with different thicknesses and preparation protocols including FFPE thin sections, fixed thick mouse brain sections cut by vibratome, and fresh thick mouse brain sections cut by hand. The results show that the CycleGAN model is capable of transforming MUSE images into Deep-MUSE images that are comparable to the brightfield H&E-stained images regardless of the sample thickness.

2.5. Evaluation metrics

To evaluate the Deep-MUSE image quality of different mouse brain sample, we use three quantitative metrics including maximum mean discrepancy (MMD), Pearson correlation coefficient, and peak signal-to-noise ratio (PSNR).

The simplified mathematic form of MMD can be represented as:

$$\text{MMD}[F, X, Y] = \left\| \frac{1}{n} \sum_{i=1}^n f(x_i) - \frac{1}{m} \sum_{j=1}^m f(y_j) \right\|_{\mathcal{H}} \quad (14)$$

where image data distribution X has n samples, image data distribution Y has m samples, and the \mathcal{H} term refers to kernel embedding of distributions.

Pearson correlation coefficient is defined as:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (15)$$

where $\text{cov}(X, Y)$ stands for covariance between data distributions X and Y , σ_X and σ_Y refer to the standard deviations of data distributions X and Y , respectively. In our case, a Deep-MUSE image x and an H&E-stained image y were used as data distributions X and Y .

PSNR is defined as:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (16)$$

where MAX_I equals to 255 for 8-bit image data, and MSE represents the mean squared error between an input image X and a target image Y with an image size of $m \times n$. For color images, the MSE of R , G , and B channels were averaged.

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [X(i, j) - Y(i, j)]^2 \quad (17)$$

3. Results

3.1. Imaging of FFPE thin mouse brain sections without and with H&E staining

To initially validate the performance of our MUSE imaging system and the associated deep learning algorithm for style transformation, several 7- μm FFPE mouse brain sections with Hoechst staining are imaged by our MUSE system. The MUSE images (Fig. 2(A), 2(E), 2(I), and 2(M)) are subsequently transformed into Deep-MUSE images by both pix2pix (Fig. 2(B), 2(F), 2(J), and 2(N)) and CycleGAN (Fig. 2(C), 2(G), 2(K), and 2(O)) models. We scanned the FFPE thin mouse brain section over a region of interest (ROI) with 7.4 mm \times 5.2 mm along the x and y axes, respectively. The scanning time was ~ 3 mins. The entire MUSE image of this section is presented in Fig. 2(A). As a result of the excited fluorescence signal from the FFPE sample with Hoechst staining, cell nuclei appear to be green and the cytoplasm is dark blue in MUSE images. After MUSE imaging, the same slice was stained with H&E, and then imaged with a digital slide scanner for comparison (Fig. S1). The cell nuclei (purple) and cytoplasm (pink) are shown in the corresponding bright-field H&E-stained image (Fig. 2(D)). Since these two image modalities

were performed on the same slice, the cell distribution and structure should match well between the MUSE images (Fig. 2(A), 2(E), 2(I), and 2(M)) and the corresponding H&E-stained images (Fig. 2(D), 2(H), 2(L), and 2(P)).

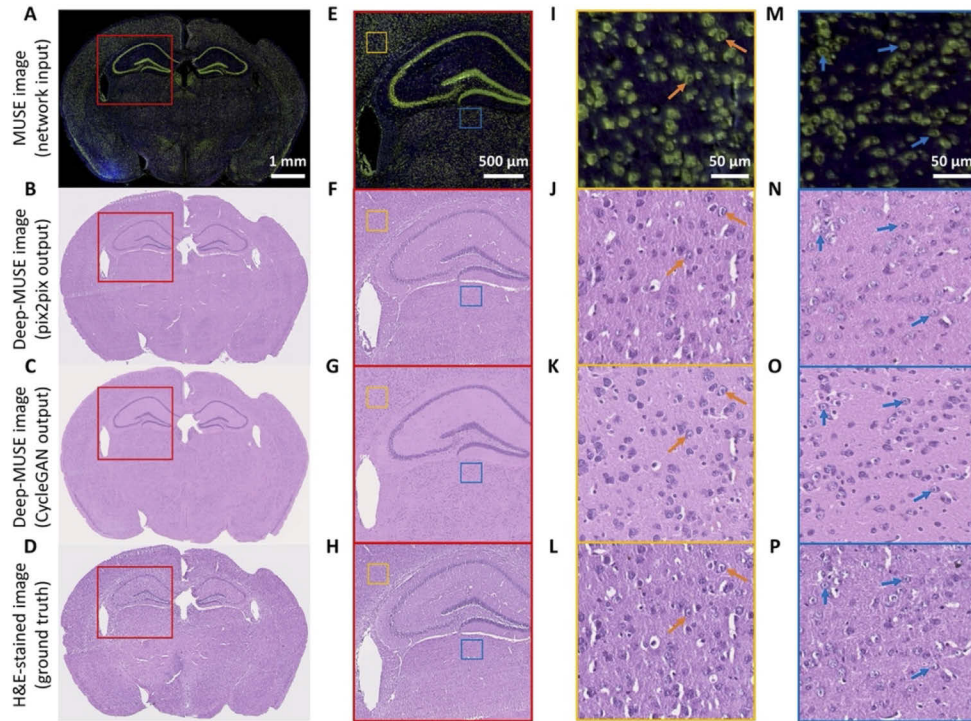


Fig. 2. Histological images of a thin mouse brain section obtained by different methods. (A) Hoechst-stained mouse brain FFPE thin slice imaged by the MUSE system. The MUSE image is used as an input into different GAN models. (B) A Deep-MUSE image output by the pix2pix model. (C) A Deep-MUSE image output by the CycleGAN model. (D) A brightfield image of the same mouse brain slice with H&E staining. (E, F, G, H) Zoomed-in images of the hippocampus (red solid regions) in (A), (B), (C), and (D), respectively. (I, J, K, L) Zoomed-in images of the yellow solid regions in (E), (F), (G), and (H), respectively. Note that all these images reveal similar cellular structures and recognizable nucleoli (orange arrows). (M, N, O, P) Zoomed-in images of the blue solid regions in (E), (F), (G), and (H), respectively. Note that all these images show comparable nucleolus structures (blue arrows).

The Deep-MUSE images are generated from the same MUSE image (Fig. 2(A)) by the pix2pix model (Fig. 2(B)) and CycleGAN model (Fig. 2(C)). Note that this MUSE image dataset was only used for validation, and the well-trained models were built on other training datasets. For pix2pix model training, the MUSE and H&E-stained image pairs of the same FFPE slice were obtained after the rigid image registration process. For CycleGAN model training, the same MUSE image for pix2pix model training was selected. However, the H&E-stained image was chosen from our digital H&E-stained image database because the CycleGAN model does not require paired data. Both GAN models learned to translate the MUSE images into Deep-MUSE images that are equivalent to the conventional H&E-stained images after training. The MUSE image (Fig. 2(E)) and the corresponding H&E-stained image (Fig. 2(H)) of the hippocampus region have a great similarity in the cell distribution. The Deep-MUSE images by the pix2pix model (Fig. 2(F)) and CycleGAN model (Fig. 2(G)) preserved the cell distribution from the MUSE image (Fig. 2(E)), and possessed the H&E color style that is comparable to the ground truth H&E-stained image

(Fig. 2(H)). The two zoomed-in regions (yellow (Fig. 2(I)) and blue (Fig. 2(M))) marked in the MUSE image (Fig. 2(E)) show subcellular features such as nucleoli, marked with orange and blue arrows. The corresponding regions (Fig. 2(L) and 2(P)) in the H&E-stained image (Fig. 2(H)) have similar features and cellular details, marked with orange and blue arrows, which indicates that the original MUSE images can provide high imaging contrast and high resolution as the conventional bright-field images with H&E staining. To further improve the interpretability and applicability of our MUSE images by pathologists, Deep-MUSE images can be used to provide a similar style as conventional H&E-stained images, which can be generated by the pix2pix model (Fig. 2(J) and 2(N)) and CycleGAN model (Fig. 2(K) and 2(O)). [Visualization 1](#) shows a series of closeup images of a column in the ROI of an FFPE mouse brain section imaged by MUSE and the corresponding H&E-stained images using different methods. The cellular structures and cytoplasm can be well identified in the closeup scanning images. The ground truth H&E-stained images (Fig. 2(L) and 2(P)) show many white cracks and suffer from information loss due to the subsequent sample preparation in the H&E staining process, e.g., washing by distilled water, dehydration using graded ethanol, vitrification by xylene. Deep-MUSE images (Fig. 2(K) and 2(O)) preserve the cellular and cytoplasm information from the original MUSE images (Fig. 2(I) and 2(M)).

To evaluate the performance of the pix2pix and CycleGAN models, we used two quantitative metrics — Pearson correlation coefficient and PSNR, to further analyze the Deep-MUSE image quality of the FFPE thin mouse brain section (Fig. 3 and Table 1). The corresponding ground truth H&E-stained images were used as the reference images in the evaluation. Pearson correlation coefficient is a statistical measure reflecting the linear correlation between two data variables. The value ranges from zero to one. The higher the value (i.e., closer to one), the higher the image similarity between the Deep-MUSE image and the ground truth H&E-stained image. PSNR is a commonly used indicator to measure image quality, especially in image super-resolution and reconstruction tasks [35]. The higher the PSNR value, the better the image quality.

Table 1. Evaluation of Deep-MUSE image quality generated by different GAN models using the FFPE thin mouse brain section

Model	Dataset type	Pearson correlation coefficient (mean \pm std)	PSNR (mean \pm std)	SSIM (mean \pm std)
pix2pix	Training dataset	0.88 \pm 0.02	23.6 \pm 0.9	0.53 \pm 0.06
	Validation dataset	0.73 \pm 0.05	19.5 \pm 1.1	0.27 \pm 0.08
CycleGAN	Training dataset	0.70 \pm 0.07	19.5 \pm 1.4	0.30 \pm 0.09
	Validation dataset	0.71 \pm 0.07	19.0 \pm 1.3	0.37 \pm 0.08

For the training dataset, the distributions of Pearson correlation coefficient and PSNR show that the quality of the Deep-MUSE images generated by the pix2pix model is higher than that of the Deep-MUSE images generated by the CycleGAN model (Fig. 3(A) and 3(C)). However, we should consider the performance of the models on the validation dataset for practical applications. The comparison of the training dataset is not favorable to the CycleGAN model because the pix2pix model learns directly from the ground truth H&E-stained image data with paired training. In practice, the quality of Deep-MUSE images generated by the pix2pix model drops significantly on the validation dataset, which becomes comparable to the Deep-MUSE image quality generated by the CycleGAN model (Fig. 3(B) and 3(D)). Table 1 shows the detailed statistical values with different models and datasets. The pix2pix model obtained higher values in the Pearson correlation coefficient, PSNR, and SSIM than that of the CycleGAN model on the training dataset. However, for the validation dataset, the Deep-MUSE images generated by the two models have nearly identical image quality with similar values of Pearson correlation coefficient and PSNR. The SSIM of Deep-MUSE images on the validation dataset shows that the CycleGAN model has

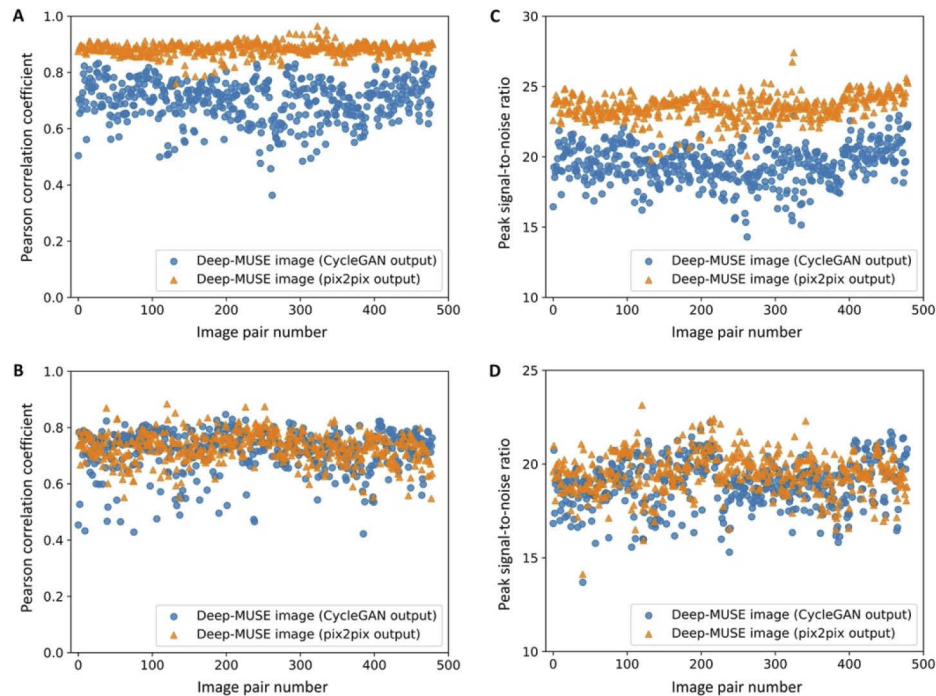


Fig. 3. Image quality comparison of different Deep-MUSE images generated by pix2pix and CycleGAN models. (A) Pearson correlation coefficient between Deep-MUSE and ground truth H&E-stained image pairs tested on the training dataset. (B) Pearson correlation coefficient between Deep-MUSE and ground truth H&E-stained image pairs tested on the validation dataset. (C) PSNR between Deep-MUSE and ground truth H&E-stained image pairs tested on the training dataset. (D) PSNR between Deep-MUSE and ground truth H&E-stained image pairs tested on the validation dataset.

better generalization than the pix2pix model with a higher value. These qualitative evaluation metrics prove that the Deep-MUSE image generated by the CycleGAN model is stable on both training and validation datasets.

3.2. Deep-MUSE image of a fixed thick mouse brain and liver versus conventional H&E-stained image

To demonstrate the possibility of the proposed histopathology workflow using Deep-MUSE images with the CycleGAN model for thick tissue, we imaged a formalin-fixed thick mouse brain and liver tissues. The MUSE image was then transformed into a Deep-MUSE image using the well-trained CycleGAN model with minimal and stable loss (Fig. S6). Subsequently, the imaged sample was further processed following the standard histopathology workflow to get the H&E-stained image of the superficial layer as a reference. Note that the thick specimen surface imaged by the MUSE system is not exactly the reference H&E-stained thin slice due to the unavoidable difference in orientation and imaging thickness. Therefore, only the CycleGAN model can be utilized for a thick tissue style transformation, further demonstrating the importance of unsupervised learning in this application.

We have imaged a fixed mouse brain specimen. The scanning time was ~7 mins to image the entire cross-section with an area of 10.9 mm × 7.0 mm (Fig. 4(A)). Figure 4 shows this validation dataset using the well-trained CycleGAN model. The Deep-MUSE image generated by the

CycleGAN model (Fig. 4(B)) and the reference H&E-stained image (Fig. 4(C)) provide similar cellular features, which can be easily identified in the hippocampus regions. The images of the left hippocampus region (Fig. 4(D) to 4(F), red regions) and right hippocampus region (Fig. 4(G) to 4(I), purple regions) show similar distributions of cell nuclei. The yellow (Fig. 4(J) to 4(L)) and blue regions (Fig. 4(M) to 4(O)) show the individual cell nuclei including their nucleoli. Both the Deep-MUSE images (Fig. 4(K) and 4(N)) and original MUSE images (Fig. 4(J) and 4(M)) can distinguish nucleoli from dark purple color and bright green color, respectively, which are marked with orange arrows (Fig. 4(J) and 4(K)) and blue arrows (Fig. 4(M) and 4(N)). [Visualization 2](#) shows a series of closeup MUSE images of a column of the ROI, the corresponding Deep-MUSE images, and the H&E-stained images of the adjacent layer. The style and the cellular information of Deep-MUSE images generated by the CycleGAN model are comparable to the conventional H&E-stained images.

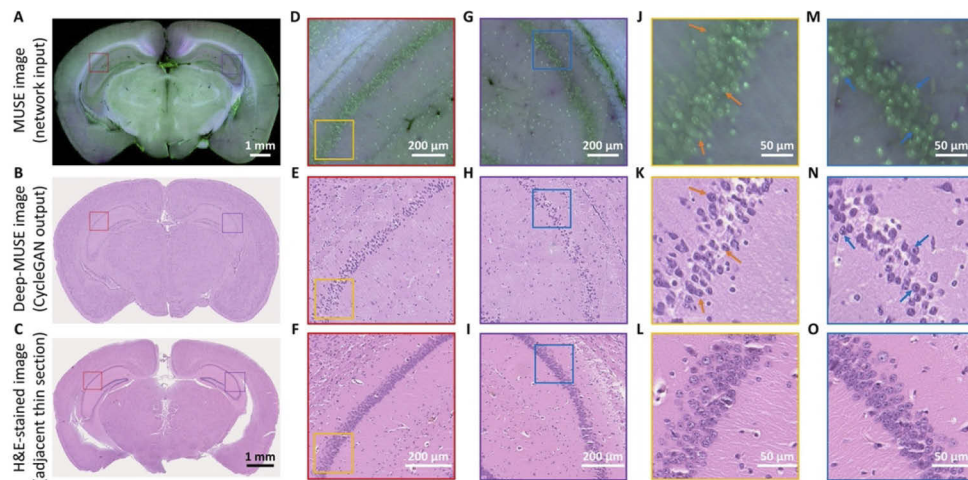


Fig. 4. Comparison between MUSE and Deep-MUSE images of a fixed thick mouse brain with H&E-stained images of the adjacent thin section. (A) MUSE image of a thick mouse brain stained by Hoechst. (B) Deep-MUSE image generated by CycleGAN. (C) Reference H&E-stained image of the adjacent thin section shown in (A) obtained after paraffin embedding, sectioning, and H&E-staining. (D, E, F) Zoomed-in images of the left hippocampus (solid red regions) in (A), (B), and (C), respectively. (G, H, I) Zoomed-in images of the right hippocampus (solid purple regions) in (A), (B), and (C), respectively. (J, K, L) Zoomed-in images of the yellow regions in (D), (E), and (F), respectively. Note that similar cellular structures (orange arrows) can be observed in (J) and (K). (M, N, O) Zoomed-in images of the blue regions in (G), (H), and (I), respectively. Note that similar cellular structures (blue arrows) can be observed in (M) and (N).

We also imaged a fixed and thick mouse liver specimen to explore the applicability of Deep-MUSE on other organs. Figure 5 shows MUSE images (Fig. 5(A) and (B)), Deep-MUSE images (Fig. 5(C) and (D)), and reference H&E-stained images (Fig. 5(E) and (F)) from two regions of the same mouse liver sample. The Deep-MUSE images (Fig. 5(C) and (D)) generated by the CycleGAN model and original MUSE images (Fig. 5(A) and (B)) have similar structures such as hepatocytes and blood vessels marked with green and orange arrows, respectively. These common cellular features are comparable in the reference H&E-stained images (Fig. 5(E) and (F)). The results demonstrate that Deep-MUSE imaging can also be applied to other tissues and organs.

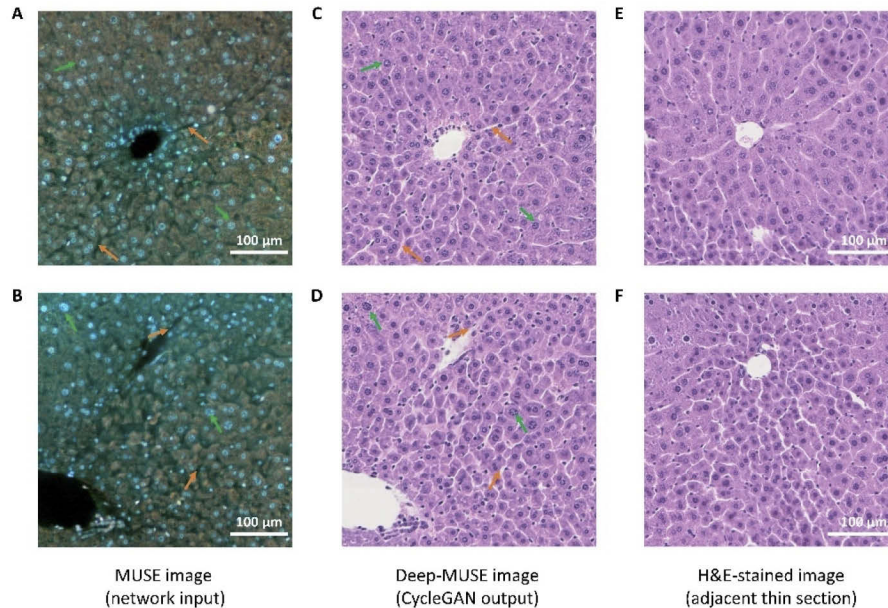


Fig. 5. Comparison between MUSE and Deep-MUSE images of a fixed and thick mouse liver with H&E-stained images of the adjacent thin section. (A, B) MUSE images of a thick mouse liver stained by DAPI. Note that hepatocytes and blood vessels are marked with green and orange arrows, respectively. (C, D) Deep-MUSE images generated by the CycleGAN. Note that hepatocytes and blood vessels corresponding to (A) and (B) are marked with green and orange arrows, respectively. (E, F) Reference H&E-stained images of the adjacent thin section.

3.3. Deep-MUSE image of a fresh thick mouse brain versus traditional H&E-stained image

To illustrate the full potential of our novel approach for clinical and intraoperative applications, our proposed histopathology workflow using Deep-MUSE images with the CycleGAN model has been applied to fresh mouse brain specimens. Similarly, the bright-field H&E-stained image of the adjacent thin section was obtained for validation.

For the fresh mouse brain specimen, the ROI was 10.7 mm × 7.3 mm, and the scanning time was ~6.5 mins. Figure 6 shows this validation dataset using the well-trained CycleGAN model. Note that the color tone of the fresh thick mouse brain image (Fig. 6(A)) is different from the fixed thick mouse brain image (Fig. 4(A)) due to different biomolecular properties with/without fixation and different white balance ratio of the camera. Therefore, the parameters of the CycleGAN model are unique for samples with different preparation protocols. The models were trained separately. The MUSE (Fig. 6(A)) and Deep-MUSE images (Fig. 6(B)) show identical features in the hippocampus region. The red (Fig. 6(C)) and yellow regions (Fig. 6(F)) are from the right hippocampus in the MUSE image (Fig. 6(A)). The corresponding Deep-MUSE images output by the CycleGAN model (Fig. 6(D) and 6(G)) and the reference H&E-stained images (Fig. 6(E) and 6(H)) show similar cellular features and distributions. Individual cells can be observed in the original MUSE images (Fig. 6(C) and 6(F)), the Deep-MUSE images (Fig. 6(D) and 6(G)), and the reference H&E-stained images (Fig. 6(E) and 6(H)).

To evaluate the Deep-MUSE image quality, we adopted qualitative and quantitative methods for both fixed or fresh thick tissue samples. First, we used the pre-trained deep neural network VGG16 [36] to extract the 4096 dimension feature vector of the input images for further analysis,

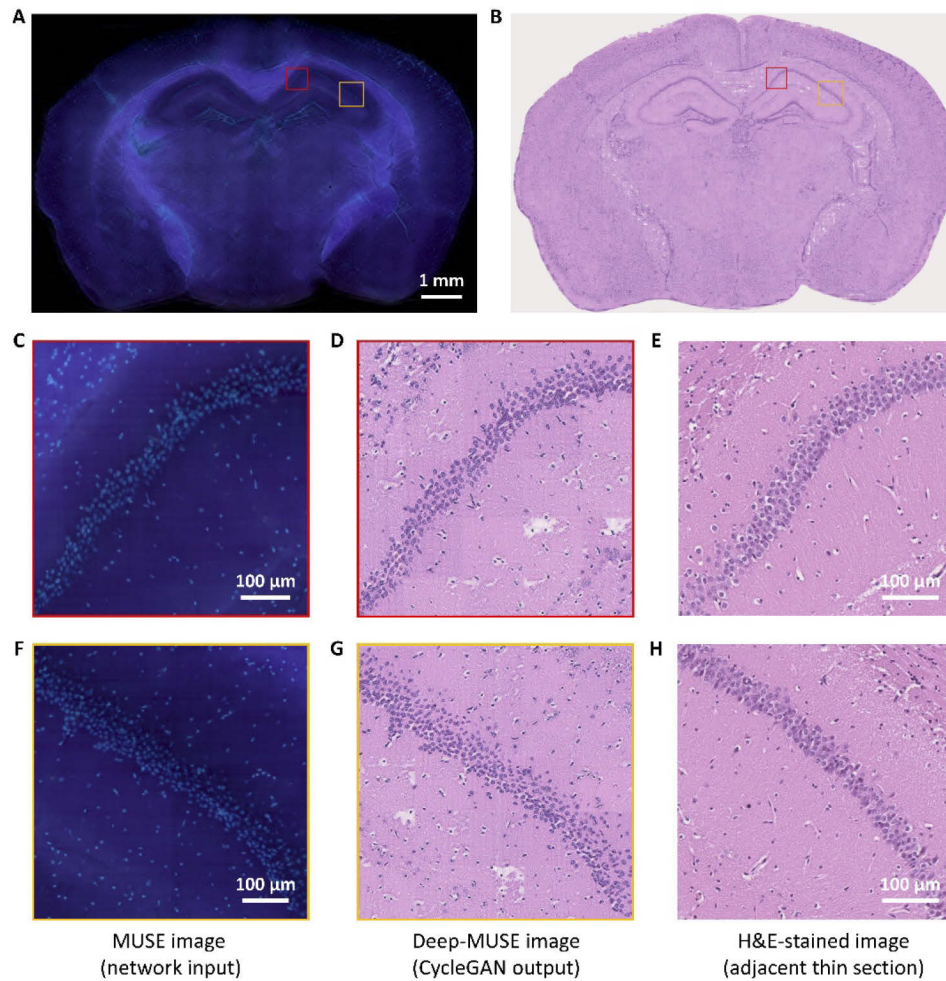


Fig. 6. Comparison between MUSE and Deep-MUSE images of a fresh mouse brain with H&E-stained images of the adjacent thin section. (A) MUSE image of a fresh mouse brain stained by Hoechst. (B) Deep-MUSE image output generated by CycleGAN. (C and D) Zoomed-in MUSE and Deep-MUSE images of the red regions in (A) and (B), respectively. (F and G) Zoomed-in MUSE and Deep-MUSE images of the yellow regions in (A) and (B), respectively. (E and H) H&E-stained images of the adjacent thin section from the same sample, which are corresponding to (C) and (F), respectively.

which can reduce data volume and reserve the discriminative information. For qualitative measurement, we applied t-SNE [37] to visualize the data on a two-dimensional graph. We chose three image datasets for visualization, which include MUSE images, H&E-stained images for model training, and Deep-MUSE images generated by the model. The visualization diagrams for the fixed mouse brain specimen (Fig. 7(A)) and fresh mouse brain specimen (Fig. 7(B)) show that the clear separation of the MUSE images from the Deep-MUSE and the H&E-stained images. The Deep-MUSE images mixed with the conventional H&E-stained images, which indicates that the Deep-MUSE images generated by the deep learning model mimic the style of the specified H&E-stained images well.

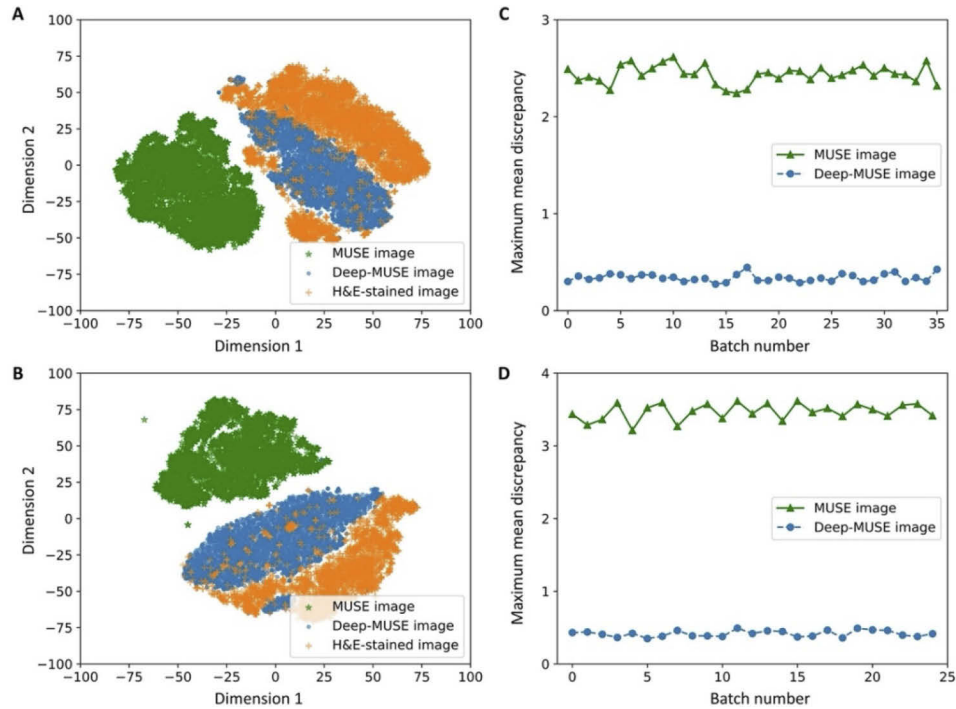


Fig. 7. Evaluation of MUSE and Deep-MUSE image quality for thick mouse brain specimens. (A) T-SNE visualization of MUSE, Deep MUSE, and H&E-stained images for the fixed mouse brain. (B) T-SNE visualization of MUSE, Deep MUSE, and H&E-stained images for the fresh mouse brain. (C) The MMD distance from MUSE and Deep-MUSE images to the H&E-stained images for the fixed mouse brain. (D) The MMD distance from MUSE and Deep-MUSE images to the H&E-stained images for the fresh mouse brain.

To analyze the Deep-MUSE image quality quantitatively, we calculated the maximum mean discrepancy (MMD) [38] to measure the distance between the generated image distribution and the target image distribution [39]. In our case, we computed the MMD distance from Deep-MUSE image features to conventional H&E-stained image features. To provide a reference, the MMD distance from the original MUSE image features to the conventional H&E-stained image features was also computed for comparison. The value of MMD is based on every batch with 100 images sampled from each data distribution. The MMD distance for the fixed mouse brain specimen (Fig. 7(C)) and fresh mouse brain specimen (Fig. 7(D)) reveals a similar conclusion that the distance from the generated Deep-MUSE images to the conventional H&E-stained images is much shorter than the distance from the original MUSE images to the conventional H&E-stained images. This result is in line with our expectation as the CycleGAN model learns the complex

style mapping function between the MUSE images and H&E-stained images during the training stage. The Deep-MUSE images output by our model is close to the conventional H&E-stained images. Table 2 shows the statistical data of MMD distances for both fixed and fresh thick mouse brain samples, corresponding to the Fig. 7(C) and 7(D).

Table 2. MMD distances of MUSE and Deep-MUSE images to training H&E-stained images using different thick mouse brain samples

Tissue type	Image category	Value (mean \pm std)
Fixed mouse brain	MUSE image	2.436 ± 0.092
	Deep-MUSE image	0.338 ± 0.039
Fresh mouse brain	MUSE image	3.470 ± 0.113
	Deep-MUSE image	0.417 ± 0.042

4. Discussion

We demonstrated that the MUSE imaging system integrated with a deep learning algorithm is able to provide rapid slide-free histological images (i.e., Deep-MUSE) which are equivalent to standard H&E-stained images. This data-driven approach has great potential to become an alternative intraoperative histopathology workflow, replacing the needs of any sample processing (except staining) while providing high-quality images that are close to FFPE thin slices with H&E staining. To prove the reliability of our proposed workflow, we compared the performance of the CycleGAN model and the state-of-the-art pix2pix model on FFPE thin slices with the ground truth H&E-stained images as the reference (Fig. 2 and Visualization 1). Afterward, the presented workflow was tested on unsectioned fixed mouse brain and liver specimens to further explore the rapid and slide-free advantages of the MUSE imaging technique (Fig. 4, Fig. 5 and Visualization 2). Finally, we directly apply our workflow to fresh mouse brain samples (Fig. 6). The correspondence between Deep-MUSE images and standard H&E-stained images shows that our approach can assist pathologists for histopathological examination.

The superiority of the proposed histopathology workflow with Deep-MUSE images can be elaborated in two aspects. First, it is rapid to acquire Deep-MUSE images with similar molecular contrast compared with the conventional H&E-stained images. For the MUSE imaging system, its slide-free imaging capability allows imaging of thick tissue directly with UV excitation light, bypassing time-consuming sample preparation steps that are needed to obtain optically thin sections in standard histopathology workflow. It will only take ~ 30 seconds for hand cutting, ~ 2 mins for Hoechst staining, and ~ 2 min for acquiring a MUSE image with subcellular resolution for a typical brain biopsy ($5 \text{ mm} \times 5 \text{ mm}$). As for training the deep neural network, dedicated hardware (e.g., graphics processing units (GPUs)) can be used for acceleration. For example, the model training of the fresh mouse brain network accelerated by a single GPU requires only 20 hours. More importantly, once the model obtains optimal parameters after training, the feed-forward network outputs the Deep-MUSE images rapidly with the speed of 1.6 s mm^{-2} by a single GPU acceleration. It will take ~ 40 seconds to generate a $5 \text{ mm} \times 5 \text{ mm}$ Deep-MUSE image from a $5 \text{ mm} \times 5 \text{ mm}$ MUSE image. Therefore, our approach will require ~ 6 mins in total to obtain a $5 \text{ mm} \times 5 \text{ mm}$ Deep-MUSE image that is comparable to a standard H&E-stained image. With the proposed Deep-MUSE workflow, it is possible to provide rapid histology ~ 6 mins. Second, the deep learning model outputs virtual H&E-stained images with a consistent style, and reduces the color variations influenced by different lab workers and staining conditions, thus providing a unified color tone for accurate diagnoses. We compared the performance of pix2pix and CycleGAN models, finding that the CycleGAN model provides images as high quality as the pix2pix model, while offering higher flexibility. For instance, for FFPE mouse brain thin sections, the CycleGAN model has comparable performance with the pix2pix model

on the validation dataset. However, the pix2pix model requires paired MUSE and H&E-stained images with a high correlation for training. The correlation between these image pairs will be easily affected by the physical staining process and image registration quality. On the contrary, the H&E-stained image data for CycleGAN model training does not have such strict requirements due to the unpaired training mode. For thick mouse brain samples, the pix2pix model cannot be applied as there is no paired image data for the model training. It is also not possible to directly test MUSE images of the thick tissue on a well-trained pix2pix model trained by FFPE thin section because the color tone of the MUSE images of thin section and thick tissue is different (Fig. S7). The CycleGAN model can easily select H&E-stained images from the database for model training, which simplifies the data pre-processing procedures by eliminating the need for image registration.

The limitations of the proposed workflow can be stated in the following ways. First, three models corresponding to mouse brain samples prepared by three different sample preparation protocols must be trained separately because of the color tone variation (Fig. 2(A), 4(A), and 6(A)). One possible solution is to modify our system with a monochrome camera to acquire grayscale images. The unified CycleGAN model can be trained with the mixed grayscale image data of thin and thick tissue samples. Second, our current model can only be applied to obtain virtual H&E staining because other staining methods, like Masson's trichrome stain and Jones' silver stain, have not been added to the training dataset. To achieve the multi-staining network and extend the application scenarios, Deep-MUSE images with individual staining or a mixture of multiple staining can be generated by simply adding the condition label to different stained image domains, e.g., a digital staining matrix to the input image for paired training [40], or domain labels for unpaired multi-domain image-to-image translation [41]. Finally, we have focused on normal mouse brain and liver tissues in this study. More investigations are needed to test for the style transformation accuracy on human specimens, especially for cancerous tissues from different organs.

The potential applications for the presented workflow are clear. Deep-MUSE images provide high molecular contrast with subcellular resolution without laborious sample preparation (e.g., staining only), which can enable intraoperative pathologic assessment. For example, during a surgical operation, a fresh human tissue excised by a surgeon in the operating room can be placed into our system after staining. Subsequently, a high-resolution Deep-MUSE image for a typical brain biopsy (5 mm × 5 mm) will be obtained ~6 mins, which can provide an alternative to intraoperative frozen section histology which takes ~30 mins. Different cells in Deep-MUSE images can also be automatically classified and graded by deep learning algorithms [42–44]. We believe that our proposed novel workflow holds great promise as an accurate and intraoperative assessment tool used by surgeons and pathologists in hospitals.

Funding. The Hong Kong University of Science and Technology (R9421); Research Grants Council, University Grants Committee (26203619); The Hong Kong Government Innovation and Technology Commission (ITS/036/19).

Acknowledgments. We thank L. Kang and Y. Zhang for the meaningful discussion.

Disclosures. T.T.W.W. has a financial interest in PhoMedics Limited, which, however, did not support this work. The rest of the authors declare that they have no competing interests.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. X. H. Gao, J. Li, H. F. Gong, G. Y. Yu, P. Liu, L. Q. Hao, L. J. Liu, C. G. Bai, and W. Zhang, "Comparison of fresh frozen tissue with formalin-fixed paraffin-embedded tissue for mutation analysis using a multi-gene panel in patients with colorectal cancer," *Front. Oncol.* **10**, 1–8 (2020).
2. Y. Rivenson, K. de Haan, W. D. Wallace, and A. Ozcan, "Emerging advances to transform histopathology using virtual staining," *BME Frontiers* **2020**, 1–11 (2020).

3. D. S. Gareau, Y. Li, B. Huang, Z. Eastman, K. S. Nehal, and M. Rajadhyaksha, "Confocal mosaicing microscopy in Mohs skin excisions: feasibility of rapid surgical pathology," *J. Biomed. Opt.* **13**(5), 054001 (2008).
4. T. T. W. Wong, R. Zhang, P. Hai, C. Zhang, M. A. Pleitez, R. L. Aft, D. V. Novack, and L. V. Wang, "Fast label-free multilayered histology-like imaging of human breast cancer by photoacoustic microscopy," *Sci. Adv.* **3**(5), e1602168 (2017).
5. X. Li, L. Kang, Y. Zhang, and T. T. W. Wong, "High-speed label-free ultraviolet photoacoustic microscopy for histology-like imaging of unprocessed biological tissues," *Opt. Lett.* **45**(19), 5401 (2020).
6. Y. K. Tao, D. Shen, Y. Sheikine, O. O. Ahsen, H. H. Wang, D. B. Schmolze, N. B. Johnson, J. S. Brooker, A. E. Cable, J. L. Connolly, and J. G. Fujimoto, "Assessment of breast pathologies using nonlinear microscopy," *Proc. Natl. Acad. Sci. U. S. A.* **111**(43), 15304–15309 (2014).
7. S. You, H. Tu, E. J. Chaney, Y. Sun, Y. Zhao, A. J. Bower, Y.-Z. Liu, M. Marjanovic, S. Sinha, Y. Pu, and S. A. Boppart, "Intravital imaging by simultaneous label-free autofluorescence-muliharmonic microscopy," *Nat. Commun.* **9**(1), 2125 (2018).
8. C. Sun, C. Kao, M. Wei, S. Chia, F. X. Kärtner, A. Ivanov, and Y. Liao, "Slide-free imaging of hematoxylin-eosin stained whole-mount tissues using combined third-harmonic generation and three-photon fluorescence microscopy," *J. Biophotonics* **12**, 1–21 (2019).
9. M. Wang, D. B. Tulman, A. B. Sholl, H. Z. Kimbrell, S. H. Mandava, K. N. Elfer, S. Luethy, M. M. Maddox, W. Lai, B. R. Lee, and J. Q. Brown, "Gigapixel surface imaging of radical prostatectomy specimens for comprehensive detection of cancer-positive surgical margins using structured illumination microscopy," *Sci. Rep.* **6**(1), 27419 (2016).
10. A. K. Glaser, N. P. Reder, Y. Chen, E. F. McCarty, C. Yin, L. Wei, Y. Wang, L. D. True, and J. T. C. Liu, "Light-sheet microscopy for slide-free non-destructive pathology of large clinical specimens," *Nat. Biomed. Eng.* **1**(7), 0084 (2017).
11. D. S. Gareau, "Feasibility of digitally stained multimodal confocal mosaics to simulate histopathology," *J. Biomed. Opt.* **14**(3), 034050 (2009).
12. M. G. Giacomelli, L. Husvogt, H. Vardeh, B. E. Faulkner-Jones, J. Horneegger, J. L. Connolly, and J. G. Fujimoto, "Virtual hematoxylin and eosin transillumination microscopy using epi-fluorescence imaging," *PLoS One* **11**(8), e0159337 (2016).
13. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in *Neural Information Processing Systems*, 2672–2680 (2014).
14. Y. Rivenson, H. Wang, Z. Wei, K. de Haan, Y. Zhang, Y. Wu, H. Günaydin, J. E. Zuckerman, T. Chong, A. E. Sisk, L. M. Westbrook, W. D. Wallace, and A. Ozcan, "Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning Nat," *Biomed. Eng.* **3**(6), 466–477 (2019).
15. Y. Rivenson, T. Liu, Z. Wei, Y. Zhang, K. de Haan, and A. Ozcan, "PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning," *Light: Sci. Appl.* **8**(1), 23 (2019).
16. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2017), Vol. 2017, pp. 5967.
17. M.-Y. Liu, T. Breuel, and J. Kautz, "Unsupervised image-to-image translation networks," in *Neural Information Processing Systems* (2017).
18. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *2017 IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2017), Vol. 2017, pp. 2242–2251.
19. R. Wang, P. Song, S. Jiang, C. Yan, J. Zhu, C. Guo, Z. Bian, T. Wang, and G. Zheng, "Virtual brightfield and fluorescence staining for Fourier ptychography via unsupervised deep learning," *Opt. Lett.* **45**(19), 5405 (2020).
20. P. Pradhan, T. Meyer, M. Vieth, A. Stallmach, M. Waldner, M. Schmitt, J. Popp, and T. Bocklitz, "Computational tissue staining of non-linear multimodal imaging using supervised and unsupervised deep learning," *Biomed. Opt. Express* **12**(4), 2280 (2021).
21. M. Combalia, J. Pérez-Anker, A. García-Herrera, L. Alos, V. Vilaplana, F. Marqués, S. Puig, and J. Malvehy, "Digitally stained confocal microscopy through deep learning," *Proc. Mach. Learn. Res.* 1–9 (2019).
22. T. Abraham, A. Shaw, D. O'Connor, A. Todd, and R. Levenson, "Slide-free MUSE microscopy to H&E histology modality conversion via unpaired image-to-image translation GAN Models," (2020) arXiv:2008.08579.
23. F. Fereidouni, Z. T. Harmany, M. Tian, A. Todd, J. A. Kintner, J. D. McPherson, A. D. Borowsky, J. Bishop, M. Lechpammer, S. G. Demos, and R. Levenson, "Microscopy with ultraviolet surface excitation for rapid slide-free histology," *Nat. Biomed. Eng.* **1**(12), 957–966 (2017).
24. A. Qorbani, F. Fereidouni, R. Levenson, S. Y. Lahoubi, Z. T. Harmany, A. Todd, and M. A. Fung, "Microscopy with ultraviolet surface excitation (MUSE): A novel approach to real-time inexpensive slide-free dermatopathology," *J. Cutan Pathol* **45**(7), 498–503 (2018).
25. T. Yoshitake, M. G. Giacomelli, L. M. Quintana, H. Vardeh, L. C. Cahill, B. E. Faulkner-Jones, J. L. Connolly, D. Do, and J. G. Fujimoto, "Rapid histopathological imaging of skin and breast cancer surgical specimens using immersion microscopy with ultraviolet surface excitation," *Sci. Rep.* **8**(1), 4476 (2018).
26. Register Images with Projection Distortion Using Control Points, MATLAB & Simulink Example, (<https://www.mathworks.cn/help/images/registering-an-aerial-photo-to-an-orthophoto.html>)
27. Y. Rivenson, H. Ceylan Koydemir, H. Wang, Z. Wei, Z. Ren, H. Günaydin, Y. Zhang, Z. Göröcs, K. Liang, D. Tseng, and A. Ozcan, "Deep learning enhanced mobile-phone microscopy," *ACS Photonics* **5**(6), 2354–2364 (2018).

28. G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science* **313**(5786), 504–507 (2006).
29. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *International Conference on Medical Image Computing and Computer Assisted Intervention* 9351, 234–241 (2015).
30. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE, 2017), Vol. 2, pp. 105–114.
31. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
32. Y. Taigman, A. Polyak, and L. Wolf, "Unsupervised cross-domain image generation," in *International Conference on Learning Representations (ICLR)* (2017).
33. J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European Conference on Computer Vision (ECCV)*, 694–711 (2016).
34. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE, 2016), pp. 770–778.
35. Y. Blau, R. Mechrez, R. Timofte, T. Michaeli, and L. Zelnik-Manor, "The 2018 PIRM challenge on perceptual image super-resolution," in *European Conference on Computer Vision (ECCV Workshops)* (2019), pp. 334–355.
36. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations (ICLR)* (2015).
37. G. H. Laurens van der Maaten, "Visualizing Data using t-SNE," *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
38. A. Gretton, K. Borgwardt, M. J. Rasch, B. Scholkopf, and A. J. Smola, "A kernel method for the two-sample problem," in *Neural Information Processing Systems (NIPS)*, (2006).
39. W. Bounliphone, E. Belilovsky, M. B. Blaschko, I. Antonoglou, and A. Gretton, "A test of relative similarity for model selection in generative models," in *International Conference on Learning Representations (ICLR)* (2016).
40. Y. Zhang, K. de Haan, Y. Rivenson, J. Li, A. Delis, and A. Ozcan, "Digital synthesis of histological stains using micro-structured and multiplexed virtual staining of label-free tissue," *Light: Sci. Appl.* **9**(1), 78 (2020).
41. Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "StarGAN: unified generative adversarial networks for multi-domain image-to-image translation," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (IEEE, 2018), pp. 8789–8797.
42. N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, N. Razavian, and A. Tsirigos, "Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning," *Nat. Med.* **24**(10), 1559–1567 (2018).
43. G. Campanella, M. G. Hanna, L. Geneslaw, A. Mirafior, V. Werneck Krauss Silva, K. J. Busam, E. Brogi, V. E. Reuter, D. S. Klimstra, and T. J. Fuchs, "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images," *Nat. Med.* **25**(8), 1301–1309 (2019).
44. T. C. Hollon, B. Pandian, A. R. Adapa, E. Urias, A. V. Save, S. S. S. Khalsa, D. G. Eichberg, R. S. D'Amico, Z. U. Farooq, S. Lewis, P. D. Petridis, T. Marie, A. H. Shah, H. J. L. Garton, C. O. Maher, J. A. Heth, E. L. McKean, S. E. Sullivan, S. L. Hervey-Jumper, P. G. Patil, B. G. Thompson, O. Sagher, G. M. McKhann, R. J. Komotar, M. E. Ivan, M. Snuderl, M. L. Otten, T. D. Johnson, M. B. Sisti, J. N. Bruce, K. M. Muraszko, J. Trautman, C. W. Freudiger, P. Canoll, H. Lee, S. Camelo-Piragua, and D. A. Orringer, "Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks," *Nat. Med.* **26**(1), 52–58 (2020).